



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/608,388

Source: 1 Fwo -

Date Processed by STIC: 4/2/04

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221**

**Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

## Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: <u>10/608,388</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input type="checkbox"/> Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line <b>not exceed</b> 72 characters in length. This includes white spaces.	
3 <input checked="" type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. <b>Do not use tab codes</b> between numbers; use <b>space characters</b> , instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was <b>not</b> saved in ASCII(DOS) text, as <b>required</b> by the Sequence Rules. Please <b>ensure your subsequent submission is saved in ASCII text</b> .	
5 <input type="checkbox"/> Variable Length	Sequence(s) <input type="checkbox"/> contain n's or Xaa's representing more than one residue. <b>Per Sequence Rules, each n or Xaa can only represent a single residue.</b> Please present the <b>maximum</b> number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) <input type="checkbox"/> . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) <input type="checkbox"/> missing. If intentional, please insert the following lines for <b>each</b> skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) <b>SEQUENCE CHARACTERISTICS:</b> (Do not insert any subheadings under this heading) (xi) <b>SEQUENCE DESCRIPTION:</b> SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to <b>include</b> the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) <input type="checkbox"/> missing. If intentional, please insert the following lines for <b>each</b> skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is <b>MANDATORY</b> if n's or Xaa's are present. In <220> to <223> section, please explain location of <b>n</b> or <b>Xaa</b> , and which residue <b>n</b> or <b>Xaa</b> represents.	
10 <input checked="" type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only <b>valid</b> <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is <b>required</b> when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) <input type="checkbox"/> missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is <b>MANDATORY</b> if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n/Xaa	"n" can <b>only</b> represent a single <u>nucleotide</u> ; "Xaa" can <b>only</b> represent a single <u>amino acid</u>	



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/608,388

DATE: 04/02/2004

TIME: 11:24:57

Input Set : A:\KATO SEQUENCE LISTING.TXT

Output Set: N:\CRF4\04022004\J608388.raw

3 <110> APPLICANT: KATO, Seishi  
 4 SEKINE, Shingo  
 5 YAMAGUCHI, Tomoko  
 7 <120> TITLE OF INVENTION: HUMAN MEMBRANE ANTIGEN TM4 SUPERFAMILY PROTEIN AND DNA  
 8 ENCODING THIS PROTEIN  
 10 <130> FILE REFERENCE: 2003-0907/WMC/01791  
 12 <140> CURRENT APPLICATION NUMBER: 10/608,388  
 13 <141> CURRENT FILING DATE: 2003-06-30  
 15 <150> PRIOR APPLICATION NUMBER: 059571/1996  
 16 <151> PRIOR FILING DATE: 1996-03-15  
 18 <160> NUMBER OF SEQ ID NOS: 2  
 20 <170> SOFTWARE: PatenIn Version 2.0  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 759  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Nucleic acid  
 27 <400> SEQUENCE: 1

*pp 1-3*  
 Does Not Comply  
 Corrected Diskette Needed

*invalid response -*

*see item 10 on Error Summary Sheet*

30 atg ggc cag tgc ggc atc acc tcc tcc aag acc gtg ctg gtc ttt ctc 48  
 W--> 33 Met Gly Gln Cys Gly Ile Thr Ser Ser Lys Thr Val Leu Val Phe Leu  
 34 1 5 10 15  
 36 aac ctc atc ttc tgg ggg gca gct ggc att tta tgc tat gtg gga gcc 96  
 39 Asn Leu Ile Phe Trp Gly Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala  
 40 20 25 30  
 42 tat gtc ttc atc act tat gat gac tat gac cac ttc ttt gaa gat gtg 144  
 45 Tyr Val Phe Ile Thr Tyr Asp Asp Tyr Asp His Phe Phe Glu Asp Val  
 46 35 40 45  
 48 tac acg ctc atc cct gct gta gtg atc ata gct gta gga gcc ctg ctt 192  
 51 Tyr Thr Leu Ile Pro Ala Val Val Ile Ile Ala Val Gly Ala Leu Leu  
 52 50 55 60  
 54 ttc atc att ggg cta att ggc tgc tgt gcc aca atc cgg gaa agt cgc 240  
 57 Phe Ile Ile Gly Leu Ile Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg  
 W--> 58 65 70 75 80 80 ← misaligned nos.  
 60 tgt gga ctt gcc acg ttt gtc atc atc ctg ctc ttg gtt ttt gtc aca 288  
 63 Cys Gly Leu Ala Thr Phe Val Ile Ile Leu Leu Leu Val Phe Val Thr  
 W--> 64 85 90 95  
 66 gaa gtt gtt gta gtg gtt ttg gga tat gtt tac aga gca aag gtg gaa 336  
 69 Glu Val Val Val Val Val Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu  
 W--> 70 100 105 110  
 72 aat gag gtt gat cgc agc att cag aaa gtg tat aag acc tac aat gga 384  
 75 Asn Glu Val Asp Arg Ser Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly  
 W--> 76 115 120 125  
 78 acc aac cct gat gct gct agc cgg gct att gat tat gta cag aga cag 432  
 81 Thr Asn Pro Asp Ala Ala Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln

*see item 3 on Error Summary Sheet*

## RAW SEQUENCE LISTING

DATE: 04/02/2004

PATENT APPLICATION: US/10/608,388

TIME: 11:24:57

Input Set : A:\KATO SEQUENCE LISTING.TXT

Output Set: N:\CRF4\04022004\J608388.raw

W--> 82 130 135 140  
84 ctg cat tgt tgt gga att cac aac tac tca gac tgg gaa aat aca gat 480  
87 Leu His Cys Cys Gly Ile His Asn Tyr Ser Asp Trp Glu Asn Thr Asp  
W--> 88 145 150 155 160  
90 tgg ttc aaa gaa acc aaa aac cag agt gtc cct ctt agc tgc tgc aga 528  
W--> 93 Trp Phe Lys Glu Thr Lys Asn Gln Ser Val Pro Leu Ser Cys Cys Arg  
W--> 94 165 invalid 170 175  
96 gag act gcc agc aat tgt aat ggc agc ctg gcc cac cct tcc gac ctc 576  
99 Glu Thr Ala Ser Asn Cys Asn Gly Ser Leu Ala His Pro Ser Asp Leu  
W--> 100 <180 <185 <190 misaligned nos.  
102 tat gct gag ggg tgt gag gct cta gta gtg aag aag cta caa gaa atc 624  
105 Tyr Ala Glu Gly Cys Glu Ala Leu Val Val Lys Lys Leu Gln Glu Ile  
W--> 106 195 200 205  
108 atg atg cat gtg atc tgg gcc gca ctg gca ttt gca gct att cag ctg 672  
111 Met Met His Val Ile Trp Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu  
W--> 112 210 215 220  
114 ctg ggc atg ctg tgt gct tgc atc gtg ttg tgc aga agg agt aga gat 720  
117 Leu Gly Met Leu Cys Ala Cys Ile Val Leu Cys Arg Arg Ser Arg Asp  
W--> 118 225 230 235 240  
120 cct gct tac gag ctc ctc atc act ggc gga acc tat gca 759  
123 Pro Ala Tyr Glu Leu Leu Ile Thr Gly Gly Thr Tyr Ala  
W--> 124 245 250  
126 <210> SEQ ID NO: 2  
127 <211> LENGTH: 1722  
128 <212> TYPE: DNA  
129 <213> ORGANISM: Homo sapiens  
132 <400> SEQUENCE: 2  
134 acttgctggg gtcggggctg cgcgacggcg caggggctgc ggggagcgcc gcgcaggccg 60  
136 tgcagttcct agcgaggagg cgccgcgcc attgccctc tctcggtgag cgcagccccc 120  
138 ctctccgggc cgggccttcg cgggccaccg gcgcc atg ggc cag tgc ggc atc 173  
140 Met Gly Gln Cys Gly Ile } move under codons  
141 1 5  
143 acc tcc tcc aag acc gtg ctg gtc ttt ctc aac ctc atc ttc tgg ggg 221  
146 Thr Ser Ser Lys Thr Val Leu Val Phe Leu Asn Leu Ile Phe Trp Gly  
147 10 15 20  
149 gca gct ggc att tta tgc tat gtg gga gcc tat gtc ttc atc act tat 269  
152 Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala Tyr Val Phe Ile Thr Tyr  
153 25 30 35  
155 gat gac tat gac cac ttc ttt gaa gat gtg tac acg ctc atc cct gct 317  
158 Asp Asp Tyr Asp His Phe Phe Glu Asp Val Tyr Thr Leu Ile Pro Ala  
159 40 45 50  
161 gta gtg atc ata gct gta gga gcc ctg ctt ttc atc att ggg cta att 365  
164 Val Val Ile Ile Ala Val Gly Ala Leu Leu Phe Ile Ile Gly Leu Ile  
165 55 60 65 70  
167 ggc tgc tgt gcc aca atc cgg gaa agt cgc tgt gga ctt gcc acg ttt 413  
170 Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg Cys Gly Leu Ala Thr Phe  
171 75 80 85  
173 gtc atc atc ctg ctc ttg gtt ttt gtc aca gaa gtt gtt gta gtg gtt 461  
176 Val Ile Ile Leu Leu Leu Val Phe Val Thr Glu Val Val Val Val

## RAW SEQUENCE LISTING

DATE: 04/02/2004

PATENT APPLICATION: US/10/608,388

TIME: 11:24:57

Input Set : A:\KATO SEQUENCE LISTING.TXT

Output Set: N:\CRF4\04022004\J608388.raw

```

177          90          95          100
179 ttg gga tat gtt tac aga gca aag gtg gaa aat gag gtt gat cgc agc      509
W--> 182 Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu Asn Glu Val Asp Arg Ser
183          105          110          115
185 att cag aaa gtg tat aag acc tac aat gga acc aac cct gat gct gct      557
188 Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly Thr Asn Pro Asp Ala Ala
189          120          125          130
191 agc cgg gct att gat tat gta cag aga cag ctg cat tgt tgt gga att      605
194 Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln Leu His Cys Cys Gly Ile
195 135          140          145          150
197 cac aac tac tca gac tgg gaa aat aca gat tgg ttc aaa gaa acc aaa      653
200 His Asn Tyr Ser Asp Trp Glu Asn Thr Asp Trp Phe Lys Glu Thr Lys
201          155          160          165
203 aac cag agt gtc cct ctt agc tgc tgc aga gag act gcc agc aat tgt      701
206 Asn Gln Ser Val Pro Leu Ser Cys Cys Arg Glu Thr Ala Ser Asn Cys
207          170          175          180
209 aat ggc agc ctg gcc cac cct tcc gac ctc tat gct gag ggg tgt gag      749
212 Asn Gly Ser Leu Ala His Pro Ser Asp Leu Tyr Ala Glu Gly Cys Glu
213          185          190          195
215 gct cta gta gtg aag aag cta caa gaa atc atg atg cat gtg atc tgg      797
218 Ala Leu Val Val Lys Lys Leu Gln Glu Ile Met Met His Val Ile Trp
219          200          205          210
221 gcc gca ctg gca ttt gca gct att cag ctg ctg ggc atg ctg tgt gct      845
224 Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu Leu Gly Met Leu Cys Ala
W--> 225 215          220 220          225 225          230 230 ← misaligned nos.
O/E 227 tgc atc gtg ttg tgc aga agg agt aga gat cct gct tac gag ctc ctc      893
230 Cys Ile Val Leu Cys Arg Arg Ser Arg Asp Pro Ala Tyr Glu Leu Leu
W--> 231          235 235          240 240          245 245 ← misaligned nos.
233 atc act ggc gga acc tat gca tagttgacaa ctcaagcctg agcttt      940
235 Ile Thr Gly Gly Thr Tyr Ala
W--> 236          250 250          ← misaligned nos
238 ttggtcttgt tctgatttgg aaggtgaatt gacgaggtct gctgctgttg gcctctggag      1000
240 ttcatttagt taaagcacat gtacactggg gttggacaga gcagcttggc ttttcatgtg      1060
242 cccacctact tacctactac ctgcgacttt ctttttcctt gttctagctg actcttcatg      1120
244 cccctaagat ttttaagtacg atggtgaacg ttctaatttc agaaccaatt gcgagtcag      1180
246 tagtgtggtg gaattaaagg aggacacgag cctgcttctg ttacctccaa gtggtaacag      1240
248 gactgatgcc gaaatgtcac caggctcctt cagtcttcac agtggagaac tcttggccaa      1300
250 aggtttttgc ggggaggagg aggaaaccag ctttctgggtt aagggttaaca ccagatggtg      1360
252 cccctcattg gtgtcctttt aaaaaatatt tactgtagtc caataagata gcagctgtac      1420
254 aaaatgacta aaatagattg taggatcata tggcgtatat cttggttcat cttcaaaatc      1480
256 agagactgag ctttgaaact agtgggtttt aatcaaagtt ggctttatag gaggagtata      1540
258 atgtatgcac tactgtttta aaagaattag tgtgagtgtg tttttgtatg aatgagccca      1600
260 ttcattgtaa gtcttaagct tgttggaat aatgtgccca tgtagactag caaaatagta      1660
262 tgtagatgtg atctcagttg taaatagaaa aatctaattc aataaaactct gtatcagccc      1720
264 cc      1722

```

FGI: use lower-case letters for nucleotides.

**VERIFICATION SUMMARY**PATENT APPLICATION: **US/10/608,388**

DATE: 04/02/2004

TIME: 11:24:58

Input Set : **A:\KATO SEQUENCE LISTING.TXT**Output Set: **N:\CRF4\04022004\J608388.raw**

L:33 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:58 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:64 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:70 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:76 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:82 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:88 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:93 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:94 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:100 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:106 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:112 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:118 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:124 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:182 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2  
L:227 M:112 C: (48) String data converted to lower case,  
L:231 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2  
L:236 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:2